

DETERMINACION DE LOS CONOCIMIENTOS ESENCIALES DE QUIMICA EN LOS ESTUDIANTES DEL CICLO DIVERSIFICADO EN COSTA RICA

Virginia Sánchez Molina
Juan Ml. Esquivel Alfaro

INTRODUCCION

El presente estudio describe el proceso de confección y validación de una prueba para medir conocimientos de química, y los resultados obtenidos de su aplicación en una muestra de estudiantes de undécimo año de enseñanza secundaria en Costa Rica.

REVISION DE LITERATURA

Para la elaboración de las pruebas de conocimientos existen dos modelos de evaluación: la evaluación con referencia a normas y la evaluación con referencia a criterios (Popham y Husek 1969).

Una prueba desarrollada de acuerdo con el primer modelo puede utilizarse para describir el rendimiento de forma amplia o para seleccionar individuos para un programa educativo. Evalúa al estudiante con respecto al grupo.

Una prueba elaborada según el modelo con referencia a criterios sirve para diagnosticar los conocimientos específicos que posee un individuo o la eficacia de un programa educativo específico (Wardrop, et al 1982).

Las diferencias en el uso de las pruebas de acuerdo con el modelo bajo el cual se desarrollen obedecen a distinciones esenciales que existen entre ellas y que se pueden resumir en:

- a. la forma en que se define el dominio de conductas por medir y
- b. la interpretación de los puntajes de la prueba (Hambleton, et al, 1978; Greco, 1974, Shepard, 1979; Milman, 1980).

Estos dos modelos presentan diferencias en los aspectos técnicos tales como: Validez, confiabili-

dad, especificación del dominio, longitud de la prueba, puntaje de corte, etc.

La prueba estructurada con referencia a normas, es relativamente más barata, de más larga vida y cumple un propósito general. Está basada en el dominio de un contenido específico, muestreado apropiadamente y generalmente con un ítem por objetivo. Es general y amplia. Se aplica para determinar la efectividad de un programa educativo en el logro de objetivos instruccionales generales.

La prueba desarrollada con referencia a criterios es más específica y limitada y tiene más ítems por objetivo. Se utiliza para determinar la efectividad de un programa educativo en el logro de objetivos instruccionales específicos.

Ambos tipos de prueba hacen una cuidadosa definición de los objetivos curriculares o de contenido que van a ser cubiertos, aunque la especificidad en las pruebas con referencia a criterios es mayor.

En ambos es importante la validez de contenido. La validación de contenido, en el modelo con referencia a normas es esencialmente un juicio, según Kerlinger (1975) y está dada por la pregunta ¿mide la prueba lo que se quiere medir?

Generalmente el procedimiento seguido es solicitar a jueces competentes que emitan sus juicios. Una prueba debe ser confiable, de modo que mida las cantidades "reales" de los atributos que posee el individuo (Kerlinger 1975).

Para determinar la confiabilidad se pueden emplear tres paradigmas: estabilidad, equivalencia y consistencia interna.

Como un procedimiento alternativo, que requiere los resultados de la administración de la prueba en una sola ocasión, es el llamado "método de las mitades". La prueba se divide en dos mitades equivalentes, los puntajes resultantes se correlacionan y luego se modifica el resultado por la fórmula de Spermen-Brown (que corrige la longitud) (Van Dalen y Meyer, 1978).

PROCEDIMIENTO

El procedimiento seguido para la confección de la prueba se basa en lo establecido por Thorndike y Hagen (1977) y Brown (1976).

Para la elaboración de la prueba se efectuó una lista de objetivos y contenidos de química que se deben desarrollar en el programa de Química de X año del Ciclo Diversificado.

Luego se solicitó a profesores de Química de enseñanza media y a docentes de Química de la Universidad que priorizaran y seleccionaran de esa lista los veinte objetivos que ellos consideraran que representaban los conocimientos esenciales que un estudiante de Química de X año del Ciclo Diversificado debe poseer. Se hizo un promedio aritmético de las prioridades y se establecieron los veinte objetivos más importantes. Se elaboraron ítemes de selección múltiple con cuatro alternativas, que midieran cada uno de los objetivos seleccionados, con el fin de lograr mayor objetividad y facilitar la aplicación y calificación de la prueba.

Debido a limitaciones de tiempo y a que la prueba debía ser resuelta en un máximo de 35 minutos, se decidió que los objetivos fueran medidos cada uno por un solo ítem.

La prueba construida según las condiciones fue sometida a un jurado de cuatro profesores de enseñanza media y cuatro docentes universitarios para establecer la validez de contenido y la claridad de los ítemes. Para esto se anotaron los veinte objetivos en una hoja y se indicó a la par de cada uno el número del ítem que lo medía. Esta hoja junto con la prueba y una carta explicativa se envió a los ocho jueces y se les solicitó que indicaran si los objetivos eran medidos por los ítemes señalados. Una vez obtenida esta información se procedió a analizarla para eliminar o modificar aquellos ítemes que, de acuerdo con el criterio de tres o más jueces, no midieran los objetivos que pretendían alcanzar. Fue necesario reemplazar tres ítemes. Los nuevos ítemes siguieron el mismo procedimiento de validación.

La coincidencia del juicio de los expertos se tomó como evidencia de la validez de contenido.

La prueba se aplicó a varios grupos de décimo año, tal y como estaban constituidos en el colegio, con el propósito de determinar el tiempo de la prueba, la claridad de los ítemes, de las instrucciones y la obtención de una medida de confiabilidad y consistencia interna. Para estimar la confiabilidad se aplicó el método de las mitades, de acuerdo

con Thorndike y Hagen (1977). La consistencia interna se estimó con la fórmula de Kuder-Richardson 20. Ambas dieron un valor de 0,58.

La prueba fue aplicada a una muestra de 541 estudiantes de undécimo año matriculados en 43 colegios, distribuidos proporcionalmente entre las siete regiones educativas del país, que representan el 20% de las instituciones oficiales que cuentan en su plan de estudios con la Educación Diversificada.

RESULTADOS

Los resultados de la aplicación de la prueba describen el rendimiento de los alumnos y se muestran en los cuadros 1, 2 y 3. En el cuadro No. 1 se presentan los promedios y desviaciones estándar por modalidad de colegio. En general los resultados son bajos si se comparan con el total de puntos (22 puntos).

CUADRO No. 1

Medias aritméticas y desviaciones estándar para la prueba de conocimientos de química por modalidad de colegio

Modalidad colegio	N	\bar{X}	S
Académico	321	7,3	2,47
Nocturno	100	6,1	2,23
Técnico	120	5,9	2,27
TOTAL	541	6,8	2,47

En el cuadro No. 2, se presentan los promedios y desviaciones estándar por región educativa. El rango de promedio va de 6,1 en la Región Oriental a 7,1 en la Región Central. Del total de 22 puntos de la prueba se obtuvo un máximo de 15 puntos en la Región Central y un máximo de 10 puntos en la Región Brunca.

En el cuadro No. 3 se presentan los promedios y las desviaciones estándar por sexo.

Los hombres muestran mejores logros que las mujeres en la prueba.

Se realizaron comparaciones utilizando análisis de varianza por modalidad de colegio, región educativa y sexo, que se muestran en los cuadros No. 4, No. 5 y No. 6.

La comparación estadística de los resultados mostrados en los cuadros anteriores permite plantear las siguientes hipótesis.

CUADRO No. 2

Medias aritméticas y desviaciones estándar para la prueba de conocimientos de química por región educativa

Región	N	\bar{X}	S	X máx.
Central	169	7,1	2,78	15
Oriental	36	6,1	2,31	11
Occidental	123	6,9	2,46	14
Chorotega	63	6,6	2,19	12
Brunca	58	6,3	2,04	10
Atlántica	36	6,4	2,31	11
Norte	57	6,5	2,34	12
Nacional	541	6,8	2,47	15

CUADRO No. 3

Medias aritméticas y desviaciones estándar de la prueba de conocimientos de química, por sexo

Sexo	N	\bar{X}	S
Femenino	276	6,3	2,37
Masculino	265	7,1	2,53
TOTAL	541	6,8	2,48

1. No hay diferencias significativas entre los colegios de diferente modalidad en el rendimiento sobre conocimientos esenciales de química de los estudiantes del Ciclo Diversificado en Costa Rica.
2. No hay diferencias significativas entre los colegios de diferentes regiones educativas en el rendimiento sobre conocimientos esenciales de química de los estudiantes en el Ciclo Diversificado en Costa Rica.
3. No hay diferencias significativas entre los estudiantes de diferente sexo del Ciclo Diversificado en el rendimiento sobre conocimientos esenciales de Química.

Para someter a prueba las hipótesis anteriores se llevaron a cabo análisis de varianza.

Los cuadros 4, 5 y 6 muestran los resúmenes de estos análisis.

CUADRO No. 4

Análisis de varianza para la prueba de conocimientos de química por modalidad de colegio

Fuente	G.L.	S.C.	M.C.	F _{razón}	F _{crit, 0,05; 2 y 400}
Entre	2	222	111	19,34	3,02
Dentro	538	3 084	5,73	—	—
Total	540	3 306			

CUADRO No. 5

Análisis de varianza para la prueba de conocimientos de química por región educativa

Fuente	G. L.	S. C.	M. C.	F _{razón}	F _{crit, 0,05; 1 y 400}
Entre	6	59,4	9,40	1,63	2,12
Dentro	534	3 246	6,08	—	—
Total	540	3 305			

CUADRO No. 6

Análisis de varianza para la prueba de conocimientos de química por sexo

Fuente	G. L.	S. C.	M. C.	F _{razón}	F _{crit, 0,05; 1 y 400}
Entre	1	74,1	74,1	12,34	3,86
Dentro	538	3 231	6,01	—	—
Total	539	3 305			

Se encontraron diferencias significativas a un nivel α del 5% en cuanto al rendimiento sobre conocimientos esenciales de química entre los educandos de diferente sexo y entre los educandos de colegios de diferente modalidad; por lo tanto, las hipótesis 1 y 3 planteadas en el estudio se rechazan.

No se encontraron diferencias significativas entre los educandos de diferentes regiones educativas.

Con el propósito de determinar qué modalidades de colegios se encuentran diferencias, se hicieron pruebas de contraste de medias a posteriori, según el procedimiento de Tukey-H.S.D.

En el cuadro No. 7, se muestran los resultados de este análisis.

CUADRO No. 7

Contraste de medias a posteriori, según Tukey-HSD para la prueba de conocimientos de química por modalidad de colegio

\bar{X}	Modalidad	Acadé- mico	Noctur- no	Téc- nico
5,8	Técnico	1,5*	0,3	
6,1	Nocturno	1,2*	—	
7,3	Académico	—	—	

* Pares de grupos significativamente diferentes a un = 0,05.

Los colegios académicos diurnos son significativamente diferentes a los técnicos y a los nocturnos, en el rendimiento sobre conocimientos esenciales de química.

DISCUSION

Con este estudio se logra desarrollar una prueba para medir conocimientos esenciales de química que evidencia validez de contenido y cierto grado de confiabilidad. La prueba fue usada como un instrumento de diagnóstico evaluativo del nivel de conocimientos esenciales de química que poseen los estudiantes de secundaria en Costa Rica.

Los resultados obtenidos con la aplicación de la prueba muestran un rendimiento bajo en el nivel nacional.

De un total de 22 puntos de la prueba de conocimientos esenciales de Química, el promedio oscila alrededor de siete puntos. El análisis muestra que los colegios nocturnos presentan una condición académica inferior. Esto debería llevarnos a revisar la "igualdad" de oportunidades que ofrece el sistema educativo costarricense a sus alumnos. Es evidente la necesidad de intensificar los esfuerzos hechos para mejorar la formación de los estudiantes de enseñanza media, con miras no sólo a mejorar el rendimiento en Química, sino a formar ciudadanos que promuevan el aprovechamiento racional de nuestros recursos naturales, que diseñen o adapten diferentes formas de tecnología, que realicen investigación, y que puedan participar crítica y creativamente en la toma de decisiones que requiere el país para salir del subdesarrollo.

El Ministerio de Educación Pública y la Universidad Nacional, así como las otras universidades, deben aunar esfuerzos para mejorar la formación de los profesores, tanto de nuevo ingreso como los que están en servicio.

BIBLIOGRAFIA

- Brown, F. *Principle of Educational and Psychological testing*. New York, N.Y.: Renohart and Winston, 1976.
- Greco, T.H. "Is there really a difference between criterion referenced and norm-referenced measurement?" *Educational Technology*. 1974, 22-25.
- Hambleton, R.K. "On the use of cut-off scores with criterion-referenced tests in instructional settings". *Journal of Educational Measurement*. 1978, 15, 277-90.
- Kerlinger, F. *Investigación del Comportamiento. Técnicas y Metodología*. Nueva Editorial Interamericana S.A. México, 1975.
- Millman, J. "Computer-based item generation" In R.A. Berk (Ed) *Criterion-referenced measurement. The state of the art*. Baltimore, Maryland. The John Hopkins University Press, 1980.
- Popham, W.J. y Jusek, T.R. "Implications of criterion-referenced measurement". *Journal of Educational measurement*. 1969, 6, 1-9.
- Shepard, L. "Norm-referenced vs. criterion-referenced tests". *Educational Horizons*. 1979, 57, 26-32.
- Thorndike, R. y Hagen, E. *Test y Técnicas de medición en psicología y educación*. Editorial Trillas, México, 1977.
- Wardrop, J. L. Et al. "A framework for analyzing the inference structure of educational achievement tests". *Journal of Educational Measurement*, 1982, 19, 1-18.
- Van Dalen, D.B. y Meyer, W.J. *Manual de Técnicas de la Investigación Educacional*. Editorial Paidós, Buenos Aires, 1978.